

Nutzungsinformationen elektronischer Publikationssysteme: Anwenderwünsche

Ergebnisse dreier empirischer Erhebungen

Björn Mittelsdorf, Ulrich Herb

vom 17.12.2010, überarbeitet: 28.11.2012



Computer- und Medienservice der Humboldt-Universität zu Berlin
Niedersächsische Staats- und Universitätsbibliothek Göttingen
Saarländische Universitäts- und Landesbibliothek Saarbrücken
Universitätsbibliothek Stuttgart
und
Verbundzentrale des GBV

Auf Initiative und in Abstimmung mit der
Deutschen Initiative für Netzwerkinformation e. V.



Nutzungsinformationen elektronischer Publikationssysteme: Anwenderwünsche von Björn Mittelsdorf, Ulrich Herb steht unter einer [Creative Commons Namensnennung 3.0 Unported Lizenz](https://creativecommons.org/licenses/by/3.0/).

Inhalt

1. Einleitung.....	2
2. Experteninterviews mit Informationsspezialisten und Repositorynutzern	3
Ergebnisse	3
Repositoryexperte	3
Wissenschaftler	3
Suchmaschinenexperte	4
Wissenschaftlerin	5
Retrievalexperte	6
Zusammenfassung.....	7
3. Onlinebefragung der Repositorynutzer.....	10
Teilnehmerstruktur und Ergebnisse	10
Rezeptionsgewohnheiten.....	13
Publikationsgewohnheiten.....	14
Open-Access-Positionen.....	16
Funktionen in Repositorien	16
Akzeptanz von Empfehlungsdiensten	22
Rankingoptionen	23
Zusammenfassung.....	25
4. Befragung der Repositorybetreiber	27
Ergebnisse	27
Erhebung und Anzeige von Benutzerzugriffe auf den Publikationsservices	27
Erhebung und Anzeige von Benutzerverhalten auf den Publikationsservices.....	28
Erhebung und Anzeige von Dokumentzugriffen auf den Publikationsservices.....	29
Technische Lösungen, Standardisierung und Behandlung von Verzerrungen.....	29
Zusammenfassung.....	34
Bilanz	35
5. Fazit	37
Verwendung absoluter Nutzungshäufigkeiten von Dokumenten.....	37
Standardisierung	38
Einbeziehung von Kontextinformationen und Bewertungen.....	38
Recommender/ Empfehlungsdienste.....	39
Anforderungen an Recommender.....	39
Verwendung der Nutzungsinformationen durch Servicebetreiber	39
Nicht-nutzungsdatenbasierte Features.....	40

1. Einleitung

Im Rahmen des von der Deutschen Forschungsgemeinschaft DFG geförderten Projekts „Open Access Statistik“ (OA-Statistik, OAS)¹ führte die Saarländische Universitäts- und Landesbibliothek (SULB) drei Erhebungen durch, um in Erfahrung zu bringen, welche Verwendungsmöglichkeiten Nutzer und Betreiber von Publikationsservices sich für Informationen, die bei der Nutzung wissenschaftlicher elektronischer Dokumente oder Objekte auf Repositories anfallen (im folgenden kurz Nutzungsdaten), wünschen oder vorstellen können. Um möglichst reichhaltige Rückmeldungen zu erlangen, fanden die Untersuchungen mit unterschiedlichen Zielgruppen statt.

- Experteninterviews mit Informationsspezialisten und Repositorynutzern, Mai und Juni 2009
- Onlinebefragung unter Repositorynutzern, Oktober bis Dezember 2009
- Mailbefragung unter Repositorybetreibern, September und Oktober 2009

Die Ergebnisse der Experteninterviews gingen in die Gestaltung des Onlinefragebogens ein. Dieser Bericht fasst die Ergebnisse der einzelnen Untersuchungen kurz zusammen und schließt mit einer Synopse der Ergebnisse ab. Die sich teils widersprechenden Ergebnisse der Erhebungen erklären sich aus den unterschiedlichen Zielgruppen. Im Mittelpunkt der Befragung der Repositorybetreiber standen einzig nutzungsdatenbasierte Anwendungen, wohingegen die Onlinebefragung der Nutzer und die Experteninterviews einen Schritt weiter gingen und darüber hinaus auch andere Anforderungen an Repositories zu ermitteln versuchten.

OA-Statistik wird von der DFG (Deutsche Forschungsgemeinschaft) gefördert. Das Projekt wurde maßgeblich durch die DINI (Deutsche Initiative für Netzwerkinformation e.V.)² – und dort durch die Arbeitsgruppe "Elektronisches Publizieren" – initiiert.

Die Projektpartner Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB Göttingen), der Computer und Medien Service (CMS) der Humboldt-Universität zu Berlin (HU Berlin), die Saarländische Universitäts- und Landesbibliothek Saarbrücken (SULB Saarbrücken) und Universitätsbibliothek Stuttgart (UB Stuttgart) kooperieren mit nationalen und internationalen Partnern. In der zweiten Projektphase ist die Verbundzentrale des GBV (VZG) als weiterer Projektpartner hinzugekommen.

¹ <http://www.dini.de/projekte/oa-statistik/>

² <http://www.dini.de>

2. Experteninterviews mit Informationsspezialisten und Repositorynutzern

Zur Vorbereitung der Onlineumfrage über nutzungsdatenbasierte Mehrwertdiensten in Repositories führte die SULB im Mai und Juni 2009 qualitative Interviews (Experteninterviews) mit Testpersonen unterschiedlicher beruflicher Provenienz durch. Zu den Versuchspersonen zählten:

- ein Suchmaschinenexperte
- ein Retrievalexperte
- ein Repositoryexperte
- eine Wissenschaftlerin (erfahrene Repositorynutzerin)
- ein Wissenschaftler (erfahrener Repositorynutzer)

Im ersten Teil des Interviews stand allgemein eine Bewertung bekannter und wünschenswerter Repositoryfunktionalitäten im Vordergrund, im zweiten Teil wurde stärker auf nutzungsdatenbasierte Funktionalitäten fokussiert.

Ergebnisse

Repositoryexperte

Experte 1 (RepExp, Repositoryexperte) empfände es als wünschenswert, einen Text schon vor dessen formaler Publikation in einem Journal als Preprint auf einem Repository zu publizieren. Dies wird zwar durch die Technik, aber nicht durch die Praxis der meisten Repositories abgedeckt, die den Preprint erst nach formaler Publikation in einem Journal akzeptieren. Ein weiteren Mehrwert würde eine Art Autorenunterstützung sowie eine automatische Extraktion von Metadaten aus dem Volltext beim Upload darstellen. Zudem wünschte sich RepExp eine Verknüpfung der Uploadfunktion mit der SHERPA/RoMEO-Datenbank, um automatisch Verlagspolicies einzusehen und Embargofristen beim Einspielen von Dokumenten zu verwalten und anzuwenden. Weitere Desiderate waren für RepExp:

- Normierung von Daten in Repositories
- Recommender basierend auf inhaltlichen Eigenschaften und Nutzungsmustern der Dokumente (unter der Bedingung, dass sie taugliche Ergebnisse liefern)
- Anzeige der neusten Dokumente bei fachlichen Repositories
- Nahtlose Verlinkungen zwischen Repositorydokumenten und Dokumenten in anderen Informationsspeichern (Journals, E-Plattformen, andere Repositories etc.)
- Orientierung der Repositoryfunktionalitäten an der Datenbank SCOPUS³

Nach Ansicht von RepExp müssen Nutzungshäufigkeiten einzelner Dokumente im Fachkontext betrachtet werden. Allerdings vermutet er, dass man unter Umständen aus dem Verhältnis der Volltextdownloads und Abstractviews auf die Qualität des Volltexts schließen könne.

Wissenschaftler

Experte 2 (WissExp, Wissenschaftler) betonte, dass die Nutzung von Repositories niedrigschwellig möglich sein muss, dies schließt für ihn orts- und möglichst geräteunabhängigen Zugang ein. Als Hürde empfindet WissExp den Medienbruch zwischen der Nutzung elektronischer und nichtelektronischer Werke: „Wir haben so einen Schnitt, bis zu einem gewissen Jahrgang suche ich in Papier, das wird auch so bleiben, danach elektronisch, und das ist sehr lästig.“ Er fordert die

³ <http://www.scopus.com>

Digitalisierung älterer Wissenschaftsliteratur. Zudem wünscht WissExp sich qualitätsgeprüfte soziale Netzwerke mit einem „geschützten Bereich (...), um sich mit Kollegen auszutauschen, wo dann auch klar ist, da ist jetzt kein Unternehmensberater, kein sonst wie, der irgendwas abschöpfen will. Das ist auch eine Qualitätsverbesserung.“ Automatisch erstellte Ergebnisrankings zu Suchanfragen bewertet WissExp kritisch, ihm wäre es lieber, die Rankingkriterien selbst beeinflussen zu können.

Recommendersysteme werden von WissExp positiv bewertet, allerdings folgt er der Maßgabe *Klasse statt Masse*: „Das heißt hier wären eher wenige, aber qualitativ hochwertige Links sinnvoll als zu viele.“ Zusätzlich wüsste WissExp einen Service zu schätzen, der ihm relevante Neuerscheinungen empfiehlt – basierend auf einer Analyse der Dokumente, die er bislang genutzt hat. Eine Anzeige der Nutzungshäufigkeit eines Dokument wird von WissExp als sehr wichtige Feedbackkomponente erachtet: „Weil das sozusagen, für mich ist das ja automatisiertes Feedback, mit der Chance, mich zu verbessern. Das würde das auch attraktiver noch machen, das Medium für mich.“ Zudem spekuliert WissExp (genau wie RepExp) man könne gegebenenfalls aus dem Verhältnis Abstract-View/Volltext-View die Attraktivität des Volltextes näherungsweise einschätzen: „Wenn ich jetzt ein Buch schreibe oder einen Artikel, dann kriegen Sie relativ wenig Rückmeldung. Und wenn ich jetzt aber zum Beispiel einen Artikel schreibe, und sehe, viele lesen das Abstract, aber laden es dann nicht herunter, von der Quote, kann ich mir dann denken, vielleicht ist das dann zu langweilig geschrieben.“ Allerdings relativiert WissExp ein niedriges Verhältnis von Volltextdownload zu Abstractview könnte auch auf einen uninteressanten Abstract zurückzuführen sein.

Suchmaschinenexperte

Experte 3 (SeExp, Suchmaschinenexperte) empfindet Recommenderdienste, die auf dem Lese- und Bewertungsverhalten anderer Nutzer basieren, als attraktive Zusatzfunktionen in Repositories. Eine reine Abrufhäufigkeit, die nicht mit qualitativen Daten verbunden ist, erachtet er als weniger aussagekräftig. Zusätzlich sollten seiner Meinung nach auch *Suchräume*, also die Wege der Nutzer zum Dokument, und eine „semantische Analyse der Suchanfrage“ in den Empfehlungsmechanismus miteingehen. SeExp vermutet, dass Kontextinformationen der Dokumentnutzung die Aussagekraft von Nutzerbewertungen und statistischen Auswertungen erhöhen könnten: „Also ist das beispielsweise ein Professor, der sich diese Dokumente anschaut, und daher natürlich auch aus einem gewissen Erfahrungsschatz heraus auch natürlich diese Dokumente bewerten kann, oder ist das einfach nur ein Student, der für eine Hausarbeit ein Dokument sucht und das dementsprechend bewertet hat, das ist natürlich dann eine andere Sache.“ Zugleich relativiert er die Möglichkeit aus dem formalen Status einer Person (etwa: Professorentitel) automatisch den Expertenstatus ableiten zu können und äußert Zweifel, ob die datenschutzrechtlichen Regelungen einer solch reichhaltigen Datensammlung nicht im Wege stehen. Eine weitere Hürde macht er in der Motivation der Nutzer aus, eigene Daten preiszugeben: „Wobei ich da immer ein bisschen skeptisch bin, ob die Nutzer wirklich bereit sind, diese Sachen preiszugeben. Ich meine, wenn sie selbst sich davon einen Mehrwert erhoffen, weil sie haben ja selbst dann natürlich den Mehrwert von den Daten anderer Nutzer.“ Neben dem Beispiel einer weiteren Kontextinformation (Verweildauer beim Lesen des Dokuments) erwähnt er wie die beiden vorhergehenden Testpersonen auch das Verhältnis von Volltextdownload zu Abstractview als nützliche Zusatzinformation. Weiterhin schlägt er für Repositories ein feedbackbasiertes Rankingverfahren vor, wie es auch bei Google zur Anwendung kommt: „Google macht das ja beispielsweise so, die werten oder beziehungsweise die untersuchen ganz genau, wenn jemand sich ein Dokument auswählt und dann innerhalb eines gewissen Zeitraums wieder zurück auf die Suchseite kommt, dann gehen die davon aus, wenn das ein relativ kurzer Zeitraum war, dass dieses Dokument für den Nutzer nicht hilfreich war. Und ein analoges System

wäre dann natürlich auch eine gute Methode, um im Grunde da auch Nutzerverhalten zu analysieren. Eben daraus, wie lange jemand sich ein Dokument anschaut, ob er es sich sofort herunter lädt, oder ob er sofort wieder zurück auf die Suchseite geht. Da braucht man auch relative wenig Nutzerinteraktion aufzuzeichnen.“ Als „goldenen Weg“ sieht er allerdings die explizite Nutzerbewertung eines Dokuments und fordert generell Repositories mit Feedbackfunktionen für Nutzer aufzuwerten: „Die reine Nutzungshäufigkeit eines Dokuments sagt ja im Grunde genommen nicht sonderlich viel aus. Also das ist ja im Grunde genommen eine quantitative Aussage. Was mich als Nutzer da unbedingt interessiert ist die qualitative Aussage.“ Einen weiteren Mehrwert für Repositorynutzer sieht SeExp in der Zitatanalyse für Repositorydokumente, die zusammen mit Analysen des Contents und der Nutzungshäufigkeiten der Dokumente in Recommenderservices einfließen könnten. Als zusätzliche Datenpools der Empfehlungsdienste (außerhalb der Repositories) bringt SeExp Drittanbieter wie Zotero, Bibsonomy, LibraryThing, Amazon oder Delicious ins Spiel. Zusätzlich würden Recommendersysteme seiner Ansicht nach davon profitieren, wenn Empfehlungen anhand der Ähnlichkeit zu anderen Userprofilen oder den Verhaltensmustern (z.B. eine ähnliche Suchanfrage) anderer User gegeben würden. Suchfunktionen in Repositories würde SeExp um Personalisierungen ergänzen wollen, die auch die Basis viele der von ihm gewünschten Kontextualisierungen sein könnten. Als wichtig erachtet SeExp, dass bei Dokumentempfehlungen transparent ist, warum ihm dieses Dokument empfohlen wurde. Im Gegensatz zu WissExp findet er: „Also ich denke, da kann man gar nicht genügend Daten oder genügend andere Empfehlungen präsentieren.“ Zugleich betont er aber die Wichtigkeit des Datenschutzes.

Wissenschaftlerin

Expertin 4 (WissExpIn, Wissenschaftlerin) wünscht sich in Repositorien die Verwaltung von Suchhistorien wie sie auch in Datenbanken gebräuchlich ist. Als ebenfalls nützlich würde sie die Auswahl zwischen mehreren Rankingkriterien erachten und Personalisierungsfunktionen wie das Abspeichern von Suchabfragen, die automatische Anzeige thematisch bevorzugter Dokumente oder vorgelegter Felder beim Dokumentupload. Ebenfalls wichtig erscheint ihr die intuitive Benutzbarkeit des Systems: „Ich möchte auch nicht hundertfünfzig Optionen für irgendwas haben, wenn die mich überhaupt noch nicht interessieren. Es wäre mal schön, wenn einfach ein paar Sachen gar nicht da wären, außer ich fange an, nach denen zu suchen, weil ich irgendetwas Bestimmtes möchte.“ Hinsichtlich der Visualisierung wünscht sie sich, dass „die virtuelle Welt die echte Welt abbildet.“ Sie erläutert dies an einem Beispiel: „Wenn ich jetzt einen Artikel habe aus irgendeiner bestimmten Zeitschrift, irgendeiner bestimmten Ausgabe, dann würde ich gerne da in der Nähe die anderen Artikel haben, die auch in dieser Ausgabe sind. (...) Ja, so wie wenn ich mir die Zeitschrift aus dem Regal hole, dann habe ich ja auch die Vorherigen und die Hinteren und kann so ein bisschen abchecken, was da sonst noch dabei ist.“

Hinsichtlich der Empfehlungsdienste ist WissExpIn kritisch. Dienste, die auf Autoren- und Koautoreninformationen aufbauen, empfindet sie als wenig hilfreich und erklärt dies mit ihrer fachlichen Herkunft (Linguistik), die von Einzelautorenschaften und *wissenschaftlichen Schulen* geprägt sei: „Also gut, wenn das so ist, dann ist das wahrscheinlich sinnvoll, nur bei uns ist das einfach nicht so. Es gibt halt Schulen oder. Wenn so was natürlich ein Computer herausfinden könnte, dann wäre das natürlich toll, wenn der das Feld ordnen könnte. In irgendeiner Form.“ Ebenso kann sie sich kaum vorstellen, einen Empfehlungsdienst auf Nutzungsmusteranalyse sinnvoll nutzen zu können: „Was ich nicht will, sind irgendwelche Empfehlungen für Schallplatten, die sich mit dem gleichen Thema beschäftigen oder weiß ich nicht was. Oder wer das gelesen hat, hat auch das gelesen. Also das hat mir noch niemals irgendwas gebracht. Ich meine, man liest das ja manchmal,

weil das ja sehr gerne mal irgendwo steht oder man wird dann... Also mir hat das noch nie was gebracht, was mir der Computer empfohlen hat. Wirklich noch nie.“ Anknüpfend das Wissenschaftsschulenmodell würde sie die Abbildung sozialer Netze sowie der besagten Schulen und darauf aufbauender Empfehlungen bevorzugen: „Ich guck jetzt einen von einer bestimmten Schule oder so und dann kommen da fünf andere Namen und wo dann steht, die machen doch so etwas Ähnliches, willst du dir das nicht auch mal angucken. Wenn das stimmen würde, klar, wäre das super.“ Allerdings müsste ein solcher Service Empfehlungen mit hoher Treffergenauigkeit (90%) liefern, andernfalls könnte sie auch passenden Treffern nicht vertrauen. Skeptisch bewertet WissExpIn auch Tagging oder qualitative Empfehlungen auf Basis von Nutzerbewertungen eines Dokuments und vertraut eher der eigenen Informationskompetenz.

Retrievalexperte

Experte 5 (RetExp, Retrievalexperte) würde Mehrwertdienste auf strukturellen Elementen von Repositories (Metadaten, Hierarchien) unter Verwendung infometrischer Analysen aufbauen. Grundvoraussetzungen wären allerdings sehr gut aufbereitete Metadaten und eine ausreichende Dokumentmenge: „Wenn man im Hunderttausenddokumentebereich ist, denke ich, kann man ganz gute Dienste, Zusatzdienste oder Berechnungen mit diesen Metadaten machen.“ Beispielhaft nennt RetExp die Erstellung von Co-Autoren-Netzgraphen und darauf aufbauender Berechnungen von Zentralitätswerten. Praktische Anwendungen dieser Auswertungen wären etwa Rankingoptionen, die der Nutzer bei Bedarf zuschalten und mittels derer er ein Re-Ranking vornehmen kann. Zusätzlich schlägt RetExp interaktive Suchunterstützung (*Searchterm Recommender*) in Form eines Mappings zwischen natürlich sprachlicher Nutzereingabe und Erschließungsvokabular vor. Grundlage einer solchen Funktion wären linguistische und statistische Verfahren. Zusätzlich könnte für RetExp eine Visualisierung der Dokumentmenge eines Repositories im Stile einer Mind Map, die nicht streng hierarchisch, sondern wie ein Netzwerk aufgebaut ist, eine nützliche Mehrwertfunktion für Repositories sein. Dies könnte für thematische Subsets eines Repositorybestandes auch die Form einer Cloud Navigation annehmen, in der „man die Terminologie, die häufig in diesen Dokumenten vorkommt, (...) anordnet“. Diese kann in einer simplen Form auch eine Übersicht sein, „beispielsweise wenn das jetzt wissenschaftliche Literatur ist, die in dem System ist, dann zeigt das System gleich an, ok, das sind hier die zehn wichtigsten Zeitschriften, dazu finden Sie so und so viele Dokumente, oder vierzig Prozent der Dokumente kommen aus diesen Zeitschriften, so was. (...) Oder die wichtigsten Autoren, ganz einfache Auszählung, dass man sozusagen die Viel-Publizierer in einer Dokumentenmenge.“ RetExp fände eine solche Selbstauskunft des Repositories fair: „Jemand kommt auf das System und weiß ja noch nix und muss dann erst mal irgendwelche Begriffe sich überlegen. Nee, da kommen gleich die topics, die belegt sind in dem System und dann vielleicht auch die Zeitschriften oder die Verlage, die da drin sind.“

Bei den Retrievalfunktionen legt RetExp großen Wert auf Drill-Down-Funktionen und Verfeinerungen von Suchabfragen, etwa durch Anreicherungen mit Fachterminen und fordert für Nutzer sehr leichte und intuitive Mechanismen zur Änderung der Suchabfragen. Die Verwendung von *user generated content* in Repositories betrachtet RetExp ambivalent: Einerseits würde er den Einsatz editierbarer Klassifikationen oder Thesauri ablehnen, andererseits würde seiner Meinung nach die Möglichkeit, Dokumente zu taggen, die Erschließung bereichern. Positiv bewertet RetExp Recommenderdienste, die auf Textähnlichkeit, Dokumenterschließung (v.a. beim verwendeten kontrollierten Vokabular), Ko-Downloads, der Analyse des bisherigen Nutzungsverhaltens eines Nutzers oder der Analyse der Suchhistorien beruhen. Er fordert aber zugleich, dass das Zustandekommen eines Rankings und einer Empfehlung plausibel und nachvollziehbar ist: „wenn der Rankingalgorithmus offen ist, was ich mir

von einem wissenschaftlichen Repositorium erwarte, dann sollte das ruhig mit angezeigt werden, warum das Dokument... Ist es rein textstatistisch gerankt worden oder wurde die Suchhistorie zu wie viel Prozent miteinbezogen, solche Dinge. Mehr Transparenz und Offenheit des Systems.“ Dies schließt für RetExp auch ein, dass verschiedene Elemente/Kriterien des Rankings vom Nutzer ab- und zugeschaltet werden können. Die Anzeige von Nutzungshäufigkeiten eines Dokument erachtet RetExp als hilfreich, einer Unterscheidung der Herkunft des Nutzers nach akademischen Campus-Nutzern und nicht-akademischen Nutzern über IP-Ranges misstraut er, da zuhause arbeitende Wissenschaftler der falschen Nutzergruppe zugeordnet werden. Sein Fazit formuliert RetExp wie folgt: „Visualisierungen und statistische Mehrwertdienste sind meines Erachtens wichtige künftige Entwicklungen bei den Repositorien.“

Zusammenfassung

Die Anzeige von Nutzungshäufigkeiten einzelner Dokumente wird prinzipiell positiv (RepExp, RetExp), teils sogar sehr positiv (WissExp) bewertet, allerdings wird betont, dass der Fachkontext und andere Variablen berücksichtigt werden müssen (RepExp). Die Verwertbarkeit reiner Nutzungshäufigkeiten wird teils auch relativiert, es sei denn diese würden mit Kontextinformationen (Suchräume, semantische Analyse der Suchterme, Status der Suchenden, Verweildauer beim Lesen des Dokuments, feedbackbasierte Rankings) und qualitativen Informationen (Nutzerbewertungen der Dokumente) verknüpft (SeExp, RetExp).

Überraschenderweise mutmaßen mehrere Interviewpartner, das Verhältnis aus Abstractview und Volltextdownload könne vielleicht Aufschluss über die Qualität des Volltexts geben (RepExp, WissExp, SeExp). Einschränkend wird erwähnt, dass ein niedriges Verhältnis von Volltext-Download zu Abstract View auch auf einen uninteressanten Abstract hinweisen könnte (WissExp).

Recommender werden prinzipiell als sinnvoll erachtet (RepExp, WissExp, SeExp, RetExp), allerdings wird eine hohe Qualität der Empfehlungen erwartet (RepExp, WissExp, WissExpIn). Eine Expertin (WissExpIn) bewertet Recommender kritisch und erklärt dies mit ihrer Fachzugehörigkeit, in der eher Schulen als einzelne Autoren oder Nutzungsmuster relevante Informationen auszeichnen. Recommender, die auf sozialen Wissenschaftlernetzwerken basieren, könnten allerdings auch nach Einschätzung dieser Expertin von Nutzen sein.

Bei den zugrundeliegenden Verfahren nennen die Teilnehmer unterschiedliche Ausgestaltungen:

- Analyse von inhaltlichen Eigenschaften von Dokumenten (RepExp)
- Analyse von Nutzungsmustern (RepExp)
- Empfehlen von Neuerscheinungen nach Analyse des eigenen Leseverhaltens/ Nutzungsmusters (WissExp)
- Analyse von Nutzungsverhalten und Bewertungen der Nutzer oder andere feedbackbasierte Verfahren (SeExp)
- Analyse des Dokumentinhalts in Kombination mit Nutzungsinformationen und Zitationsdaten (SeExp)
- Analyse von Nutzerprofilen und Nutzungsverhalten (z.B. Suchanfragen) (SeExp)
- Analyse von Wissenschaftlernetzwerken (WissExpIn)
- Analyse struktureller Elemente wie Metadaten oder Hierarchien, Autoren-Netzwerken inkl. der Bestimmung von Zentralitätsmaßen (RetExp)
- Analyse der Textähnlichkeit und Dokumenterschließung (RetExp)

- Analyse von Nutzungsverhalten und Ko-Downloads (RetExp)
- Analyse der Suchhistorien (RetExp)

Zudem wird von einem Experten (SeExp) angeregt, in die Analyse der Nutzungshäufigkeiten und –muster auch Daten aus anderen Quellen als den Repositories, also von externen Diensten miteinzubeziehen. Weiterhin wird Transparenz der Empfehlungen gefordert und zwei Experten (SeExp, RetExp) wollen wissen, warum einzelne Dokumente empfohlen wurden. Anders als SeExp, der qualitative Informationen, v.a. Nutzerbewertungen von Inhalten, als „goldenen Weg“ betrachtet, ist WissExpIn genau gegenüber solchen Elementen (Nutzerbewertungen, Tagging) sehr kritisch und vertraut lieber der eigenen Einschätzung. RetExp sieht die Verwendung von User Generated Content zwiespältig und schwankt zwischen der Befürchtung fehlender Normierung und der Verheißung einer reichhaltigen Erschließung.

Nutzungsbasierte Zusatzfeatures berühren nach Meinung einiger Experten (WissExp, SeExp) sensible Bereiche des Datenschutzes, dessen Bedeutung übereinstimmend betont wird. Zudem herrscht das Bedenken solche Features könnten auch Ängste der Nutzer gegen die Verwendung Ihrer Nutzungsinformationen fördern, zugleich besteht aber auch die Zuversicht qualitative hochwertige Services könnten diese Bedenken mindern (SeExp).

Als nicht-nutzungsdatenbasierte Features wurden genannt:

- Publikation von Preprints zur Diskussion mit der Community (RepExp)
- Normierung von Daten in Repositories (RepExp)
- Easy Submission und Automatisierung (Extraktion von Metadaten aus Volltexten, automatische Prüfung von SHERPA/RoMEO Policies im Upload-Prozess) (RepExp)
- Vernetzung und nahtlose Verbindung unter Repositorydokumenten und zwischen Dokumenten in anderen Angeboten (Journals, E-Plattformen, etc.) bzw. die Bündelung zusammengehöriger Items/Texte ohne Recherchebruch zwischen Informationsangeboten (RepExp, WissExpIn)
- Anpassung des Look-and-Feels an gewohnte Rechercheumgebungen wie kommerzielle Datenbanken (Referenz SCOPUS) (RepExp, WissExpIn)
- Intuitive Nutzungsmöglichkeiten von Repositories (WissExp, WissExpIn)
- Geräteunabhängigkeit des Zugangs zu Informationen (Mobile Devices) (WissExp)
- Aufhebung des Medienbruchs zwischen elektronischer Nutzung im Repository und Printnutzung (WissExp)
- Verbindung von Repositories mit zugangsgeprüften sozialen Netzwerken für Wissenschaftler (WissExp)
- Personalisierungsoptionen (WissExp, SeExp, WissExpIn), die auch zur Optimierung der Recommenderdienste dienen können (SeExp). Genannt wurden unter anderem z.B. eigene Gewichtungsmöglichkeiten für Rankings (WissExp), Verwaltung von Suchhistorien (WissExpIn). Personalisierungsfunktionen könnten ggf. auch die oben geschilderte Differenz bezüglich der Verwendung von Nutzerbewertungen lösen: Mittels konfigurierbarer Rankings und Recommender könnten Nutzer selbst bestimmen welche Kriterien in Ranking oder Recommender einfließen (RetExp).
- Flexibles Ranking/ Umschalten zwischen verschiedenen Rankingoptionen (WissExpIn, RetExp)

- Interaktive und intuitive Suchunterstützung (RetExp), z.B. durch Mapping zwischen natürlich sprachlichem und kontrolliertem Vokabular, Drilldown-Funktionen, Anreicherung mit Fachthermen.

Eine zwar auf quantitativen Daten beruhende, aber nicht nutzungsdatenbasierte Anwendung schwebt RetExp vor, der eine Visualisierung des Repositorybestandes in Form einer Mind Map vorschlägt. Die Navigation im Repository würde nicht mit starren Hierarchien oder Klassifikationen gestützt, sondern über die Ergebnisse semantischer Analysen und deren Visualisierung. Simple Ausprägungen wären Darstellung stark vertretener Themen, Verlage oder Autoren. Zur Umsetzung der Visualisierung sind allerdings intuitive Modelle (Cloud Navigation) anstelle starrer Listen erwünscht.

3. Onlinebefragung der Repositorynutzer

Von Oktober bis Dezember 2010 führte die SULB eine Onlinebefragung von Repositorybenutzern durch. Der Fragebogen stand in deutscher und englischer Form bereit und wurde auf unterschiedliche Weise zu bewerben versucht, u.a.:

- Versand über Mailinglisten (u.a. InetBib, OpenAccess Informationsplattform, American Scientist Open Access Forum)
- Ansprache verschiedener Fachgesellschaften um Versand der Einladung zur Teilnahme (die Effizienz dieser Maßnahme ist allerdings schwer einzuschätzen, da die angesprochenen Gesellschaften keine Rückmeldung gaben)
- Knowledge Exchange Verteiler
- Verlinkung von den Websites der Projektpartner

Aufgrund der Thematik und auch der Verteilungswege muss von einer recht starken Verzerrung bei den Teilnehmern ausgegangen werden, denen insgesamt eine größere Affinität zu den Themen Open Access und elektronisches Publizieren unterstellt werden kann. Dies spiegelt sich auch in der institutionellen und beruflichen Zuordnung der teilnehmenden Personen wider.

Teilnehmerstruktur und Ergebnisse

Von den insgesamt 253 Personen, die den Fragebogen beantwortenden, waren 128 Teilnehmer (TN) weiblich und 125 männlich. Der Großteil der TN hatte eine Ausbildung im Bereich der Naturwissenschaften (53) oder Geisteswissenschaften erworben (igs. 36, davon 15 in Geschichtswissenschaft und 21 in Literatur- und Kulturwissenschaften). 27 TN nannten eine Ausbildung im Bereich der Informatik/EDV, 21 entstammten den Empirischen Humanwissenschaften. 81 und damit fast ein Drittel der TN hatten eine bibliothekarische Ausbildung.

Einige TN mit bibliothekarischer Ausbildung haben vor einer postgraduierten Bibliotheksausbildung weitere Fächer studiert:

- Informationswissenschaft (2 TN)
- Philosophie und Musik (1 TN)
- Geisteswissenschaften (2 TN)
- Sprach- und Literaturwissenschaft (1 TN)
- Informatik und Architektur (1 TN)
- Pädagogik und Erziehungswissenschaft (2 TN)

Umgekehrt gaben 5 TN an, zusätzlich zum grundständigen Studium (zweimal Naturwissenschaften, dreimal Literaturwissenschaften) eine bibliothekarische Ausbildung absolviert zu haben, so dass sich der Zahl der TN mit bibliothekarischer Bildung auf 86 erhöht.

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Rechtswissenschaften	5	2,0	2,0	2,0
	Wirtschaftswissenschaften	8	3,2	3,2	5,1
	Medizin	3	1,2	1,2	6,3
	Geschichtswissenschaften	15	5,9	5,9	12,3
	Linguistik	8	3,2	3,2	15,4
	Literatur- und Kulturwissenschaften	21	8,3	8,3	23,7
	Bibliothekswesen	81	32,0	32,0	55,7
	Empirische Humanwissenschaften	21	8,3	8,3	64,0
	Naturwissenschaften	53	20,9	20,9	85,0
	Elektronische Datenverarbeitung oder Informatik	27	10,7	10,7	95,7
	Technik	6	2,4	2,4	98,0
	Handwerk	1	,4	,4	98,4
	Dienstleistung	3	1,2	1,2	99,6
	Behörde	1	,4	,4	100,0
	Gesamt	253	100,0	100,0	

Die Prävalenz bibliothekarisch tätiger TN verstärkt sich angesichts der aktuellen Tätigkeiten der Antwortenden noch weiter: Insgesamt 118 der 253 TN (46,6%) waren zum Zeitpunkt der Befragung im Bibliotheksumfeld tätig.

In welchem thematischen Bereich sind Sie zurzeit tätig?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
Rechtswissenschaften	3	1,2	1,2	1,2
Wirtschaftswissenschaften	4	1,6	1,6	2,8
Medizin	6	2,4	2,4	5,1
Geschichtswissenschaften	5	2,0	2,0	7,1
Linguistik	6	2,4	2,4	9,5
Literatur- und Kulturwissenschaften	3	1,2	1,2	10,7
Bibliothekswesen	118	46,6	46,6	57,3
Empirische Humanwissenschaften	14	5,5	5,5	62,8
Naturwissenschaften	32	12,6	12,6	75,5
Elektronische Datenverarbeitung oder Informatik	34	13,4	13,4	88,9
Technik	5	2,0	2,0	90,9
Dienstleistung	16	6,3	6,3	97,2
Behörde	4	1,6	1,6	98,8
Industrie	3	1,2	1,2	100,0
Gesamt	253	100,0	100,0	

Rezeptionsgewohnheiten

Mehr als die Hälfte der TN (145) gab an zwischen einer und zehn Fachpublikationen pro Monat zu lesen, 72 gaben an, monatlich zwischen 11 und 20 Publikationen zu lesen, 14 gaben an zwischen 21 und 50 Publikationen pro Monat zu lesen, je 4 lasen zwischen 51 und 100 resp. über 100 Publikationen pro Monat. 14 TN machten die Angabe gar keine oder im Schnitt weniger als eine Publikation pro Monat zu lesen.

Wie viele Fachpublikationen lesen Sie durchschnittlich pro Monat?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Keine	14	5,5	5,5	5,5
	1 bis 10	145	57,3	57,3	62,8
	11 bis 20	72	28,5	28,5	91,3
	21 bis 50	14	5,5	5,5	96,8
	51 bis 100	4	1,6	1,6	98,4
	mehr als 100	4	1,6	1,6	100,0
	Gesamt	253	100,0	100,0	

10 der TN gaben an noch nie eine Open-Access-Publikation gelesen zu haben, 78 hatten eine bis zehn OA-Publikationen gelesen, 102 hatten zwischen 11 und 100 Open-Access-Werke gelesen und 63 gaben an, mehr als 100 OA-Dokumente gelesen zu haben.

Wie viele frei zugängliche Fachpublikationen haben Sie schon gelesen?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Keine	10	4,0	4,0	4,0
	1 bis 10	78	30,8	30,8	34,8
	11 bis 100	102	40,3	40,3	75,1
	mehr als 100	63	24,9	24,9	100,0
	Gesamt	253	100,0	100,0	

Publikationsgewohnheiten

Was die Publikationshäufigkeit angeht, finden sich unter den TN nur wenige Vielschreiber. Mehr als die 87,4% der TN (221) hat maximal 20 Fachpublikationen veröffentlicht.

Wie viele Fachpublikationen haben Sie bereits veröffentlicht?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
Keine	85	33,6	33,6	33,6
1 bis 10	114	45,1	45,1	78,7
11 bis 20	22	8,7	8,7	87,4
21 bis 50	17	6,7	6,7	94,1
51 bis 100	7	2,8	2,8	96,8
mehr als 100	8	3,2	3,2	100,0
Gesamt	253	100,0	100,0	

Schließt man Fälle von der Auswertung aus, die zum Befragungszeitpunkt im Bibliotheksumfeld tätig waren, verändert sich die Datenlage unwesentlich: 85,9% der verbleibenden TN haben mehr als 20 Fachpublikationen veröffentlicht.

Wie viele Fachpublikationen haben Sie bereits veröffentlicht?
[Im Bibliotheksumfeld Tätige ausgeschlossen]

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
Keine	47	34,8	34,8	34,8
1 bis 10	55	40,7	40,7	75,6
11 bis 20	14	10,4	10,4	85,9
21 bis 50	10	7,4	7,4	93,3
51 bis 100	4	3,0	3,0	96,3
mehr als 100	5	3,7	3,7	100,0
Gesamt	135	100,0	100,0	

Das geringe Publikationsaufkommen gilt in gleichem Maße für Toll-Access- und Open-Access-Journale, auch wenn die Befragten offensichtlich eine Präferenz zu Open-Access-Journalen zeigen. 113 TN publizierten 1 bis 20 Artikel in Open-Access-Journalen, 91 TN publizierten im gleichen Bereich in Toll-Access-Journalen. Insgesamt gaben 105 TN an in Toll-Access-Journalen publiziert zu haben, eine geringfügig höhere Anzahl an TN (121) gab an in Open-Access-Journalen publiziert zu haben.

Wie viele wissenschaftliche Artikel publizierten Sie ausschließlich in Subskriptionsjournals?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	None	63	24,9	37,5	37,5
	1 to 10	78	30,8	46,4	83,9
	11 to 20	13	5,1	7,7	91,7
	21 to 40	5	2,0	3,0	94,6
	More than 40	9	3,6	5,4	100,0
	Gesamt	168	66,4	100,0	
Fehlend	99	85	33,6		
Gesamt		253	100,0		

Wie viele wissenschaftliche Artikel publizierten Sie ausschließlich in Nicht-Subskriptionsjournals?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	None	47	18,6	28,0	28,0
	1 to 10	104	41,1	61,9	89,9
	11 to 20	9	3,6	5,4	95,2
	21 to 40	7	2,8	4,2	99,4
	More than 40	1	,4	,6	100,0
	Gesamt	168	66,4	100,0	
Fehlend	99	85	33,6		
Gesamt		253	100,0		

Open-Access-Positionen

181 TN (71,5%) stimmten der Forderung, öffentlich geförderte Forschungsergebnisse zusätzlich zur Publikation in einem kommerziellen Verlag Open Access (i.S. des Self-Archiving, Green Road Open Access) zu stellen, völlig zu. Nur 6,3 % der Befragten stimmten der Forderung nicht zu oder enthielten sich einer Wertung.

Halten Sie es für sinnvoll öffentlich geförderte Forschungsarbeiten zusätzlich zu kommerzieller Verwertung kostenfrei zu publizieren?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig stimme überhaupt nicht zu	1	,4	,4	,4
stimme nicht zu	4	1,6	1,6	2,0
unentschieden	11	4,3	4,3	6,3
stimme zu	56	22,1	22,1	28,5
stimme völlig zu	181	71,5	71,5	100,0
Gesamt	253	100,0	100,0	

Auch die Datenlage für diese Frage ändert sich nach Exklusion von Personen, die im Bibliotheksumfeld tätig sind, wenig: 69,6% der verbliebenen TN stimmten der Aussage völlig zu, 8,1% stimmten ihr nicht zu oder enthielten sich einer Wertung.

Halten Sie es für sinnvoll öffentlich geförderte Forschungsarbeiten zusätzlich zu kommerzieller Verwertung kostenfrei zu publizieren? [Im Bibliotheksumfeld Tätige ausgeschlossen]

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig stimme nicht zu	4	3,0	3,0	3,0
unentschieden	7	5,2	5,2	8,1
stimme zu	30	22,2	22,2	30,4
stimme völlig zu	94	69,6	69,6	100,0
Gesamt	135	100,0	100,0	

Funktionen in Repositorien

Basierend auf den Ergebnissen der Experteninterviews und den daraus gewonnenen Hinweisen wurden im Onlinefragebogen Einschätzungen zu unterschiedlichen Zusatzfunktionen in Repositories erfragt.

Positiv bewertet wurden Empfehlungsdienste, die auf dem eigenen Lese- und damit Nutzungsverhalten basieren: Nur 12,3% bewerten solche Dienste als gar nicht oder selten nützlich, wohingegen 46,7% sie als häufig oder immer nützlich einschätzen.

Wie nützlich ist es, wenn Ihnen Dokumente ausgehend von Ihrem bisherigen Leseverhalten empfohlen werden?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	gar nicht nützlich	8	3,2	3,2	3,2
	selten nützlich	23	9,1	9,1	12,3
	manchmal nützlich	104	41,1	41,1	53,4
	häufig nützlich	94	37,2	37,2	90,5
	immer nützlich	24	9,5	9,5	100,0
	Gesamt	253	100,0	100,0	

Etwas weniger positiv ist die Einschätzung der TN zu Empfehlungsdiensten, die auf einem Abgleich des eigenen Nutzungsmusters mit den Nutzungsmustern anderer Personen beruhen. 92 TN (36,4%) bewerteten solche Services als häufig oder immer nützlich.

Wie bewerten Sie einen Dienst, der Ihnen Dokumente empfiehlt, die andere Personen gelesen haben, die ansonsten die gleichen Dokumente gelesen haben wie Sie?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	gar nicht nützlich	10	4,0	4,1	4,1
	selten nützlich	31	12,3	12,7	16,7
	manchmal nützlich	112	44,3	45,7	62,4
	häufig nützlich	67	26,5	27,3	89,8
	immer nützlich	25	9,9	10,2	100,0
	Gesamt	245	96,8	100,0	
Fehlend		8	3,2		
Gesamt		253	100,0		

Auch der Nutzen von Empfehlungsdiensten, die sich an den Autoren von Dokumenten orientieren, wurde überwiegend positiv bewertet. Wenn auch vielleicht noch positivere Bewertungen zu erwarten waren, da angenommen werden kann, dass die wahrgenommene Qualität eines Dokuments oft mit dessen Autoren assoziiert ist: 100 TN (39,5%) bewerteten solche Funktionen als häufig oder immer nützlich.

Bewerten Sie den Nutzen von Empfehlungen die darauf basieren welche Autoren beteiligt sind

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	gar nicht nützlich	6	2,4	2,4	2,4
	selten nützlich	27	10,7	11,0	13,4
	manchmal nützlich	113	44,7	45,9	59,3
	häufig nützlich	63	24,9	25,6	85,0
	immer nützlich	37	14,6	15,0	100,0
	Gesamt	246	97,2	100,0	
Fehlend		7	2,8		
Gesamt		253	100,0		

Ebenfalls positiv wurde der recht pauschal erfragte Nutzen von Zugriffs-/Downloadstatistiken bewertet. Nur 18,6 % der TN bewerten den Nutzen gering (gar nicht nützlich/ selten nützlich), 47,8 % bewerteten Nutzungs-/Downloadstatistiken als häufig oder immer nützlich.

Wie schätzen Sie den Wert von Zugriffs- oder Downloadstatistiken ein?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	gar nicht nützlich	11	4,3	4,3	4,3
	selten nützlich	36	14,2	14,2	18,6
	manchmal nützlich	85	33,6	33,6	52,2
	häufig nützlich	81	32,0	32,0	84,2
	immer nützlich	40	15,8	15,8	100,0
	Gesamt	253	100,0	100,0	

Etwas weniger positives Feedback ergab die präzisere Frage nach dem Nutzen der Anzeige der Downloadhäufigkeit bei einzelnen Dokumenten, 94 TN (37,1% der TN) fanden diese Information häufig oder immer nützlich.

Bewerten Sie den Nutzen einer Anzeige der Nutzungshäufigkeit bei jedem einzelnen Dokument.

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	15	5,9	6,2	6,2
selten nützlich	46	18,2	19,0	25,2
manchmal nützlich	87	34,4	36,0	61,2
häufig nützlich	59	23,3	24,4	85,5
immer nützlich	35	13,8	14,5	100,0
Gesamt	242	95,7	100,0	
Fehlend	11	4,3		
Gesamt	253	100,0		

Überwiegend negativ bewertet wurden Funktionen, die aus quantitativen Informationen (Downloadzahlen) die Qualität von Dokumenten zu bestimmen versuchen. 52,2 % der TN bewerteten solche Funktionen als überwiegend negativ, nur 15,5 % als häufig oder immer nützlich.

Für wie nützlich halten Sie es, wenn das System aus den Abrufzahlen die Qualität der Dokumente schätzt?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	63	24,9	25,8	25,8
selten nützlich	69	27,3	28,3	54,1
manchmal nützlich	73	28,9	29,9	84,0
häufig nützlich	30	11,9	12,3	96,3
immer nützlich	9	3,6	3,7	100,0
Gesamt	244	96,4	100,0	
Fehlend	9	3,6		
Gesamt	253	100,0		

Ähnlich verhält es sich mit weiteren Kontextinformationen, aus denen man zumindest experimentell versuchen könnte, die Qualität, Relevanz oder andere positive Merkmale eines Dokuments abzuleiten. Wie etwa im Fall der Häufigkeit, mit der auf ein Dokument ein Bookmark gesetzt wurde:

Wie nützlich ist es anzuzeigen, wie oft ein Dokument als Bookmark gesetzt wurde?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	36	14,2	14,9	14,9
selten nützlich	60	23,7	24,8	39,7
manchmal nützlich	92	36,4	38,0	77,7
häufig nützlich	38	15,0	15,7	93,4
immer nützlich	16	6,3	6,6	100,0
Gesamt	242	95,7	100,0	
Fehlend	11	4,3		
Gesamt	253	100,0		

Nur 54 TN (21,3%) fanden solche Informationen nützlich. Die relativ wenig positive Einschätzung dürfte aber auch teilweise auf das Wissen der TN um die technische und logische Problematik eines solchen Vorgehens zurückzuführen sein – zumindest spiegelt sich eine entsprechende Skepsis in den Kommentaren wieder. Ähnliches gilt für die Frage, inwiefern die Messung der Verweildauer bei der Dokumentnutzung als wertvoll erachtet wird. 38 der 253 TN (15,1%) bewerteten eine solche Information als häufig oder immer nützlich, 52,6% stufen sie als selten oder gar nicht nützlich ein.

Für wie nützlich halten Sie eine Bewertung von Dokumenten, anhand der Zeit die Leser auf dem Dokument verweilen?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	73	28,9	29,9	29,9
selten nützlich	60	23,7	24,6	54,5
manchmal nützlich	73	28,9	29,9	84,4
häufig nützlich	29	11,5	11,9	96,3
immer nützlich	9	3,6	3,7	100,0
Gesamt	244	96,4	100,0	
Fehlend	9	3,6		
Gesamt	253	100,0		

Die Einschätzungen zum Nutzen von Bewertungen hielten sich die Waage: 29,3 % der TN bewerteten solche Funktionen als selten oder gar nicht nützlich, 29,6 % als häufig oder immer nützlich.

Für wie nützlich halten Sie es, wenn Leser Dokumente öffentlich auf einer Notenskala bewerten?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	22	8,7	9,0	9,0
selten nützlich	52	20,6	21,3	30,3
manchmal nützlich	95	37,5	38,9	69,3
häufig nützlich	57	22,5	23,4	92,6
immer nützlich	18	7,1	7,4	100,0
Gesamt	244	96,4	100,0	
Fehlend	9	3,6		
Gesamt	253	100,0		

Positiver bewertet wurde die Abgabe qualitativer Kommentare zu Dokumenten: Nur 16,2% der TN bewerteten diese Option als selten oder gar nicht nützlich und 46,7% bewerten sie als häufig oder immer nützlich.

Für wie nützlich halten Sie es, wenn Leser Dokumente öffentlich mit einem kurzen Text bewerten?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	11	4,3	4,5	4,5
selten nützlich	30	11,9	12,3	16,8
manchmal nützlich	85	33,6	34,8	51,6
häufig nützlich	87	34,4	35,7	87,3
immer nützlich	31	12,3	12,7	100,0
Gesamt	244	96,4	100,0	
Fehlend	9	3,6		
Gesamt	253	100,0		

Vergleichbare Einschätzungen erfuhr die Frage nach dem Nutzen von Hinweisen auf die Qualität eines Dokuments: 14,7% stuften solche Hinweise als selten oder gar nicht nützlich ein, 45,9% als häufig oder immer nützlich.

Wie schätzen Sie den Nutzen von Hinweisen auf die Qualität eines Dokuments ein?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	9	3,6	3,6	3,6
selten nützlich	28	11,1	11,1	14,7
manchmal nützlich	100	39,5	39,5	54,2
häufig nützlich	86	34,0	34,0	88,1
immer nützlich	30	11,9	11,9	100,0
Gesamt	253	100,0	100,0	

Hier betonten v.a. TN, die den Nutzen gering einschätzten (gar nicht nützlich / selten nützlich), in ihren Kommentaren den Umstand, dass Qualität aus eigener Wahrnehmung und Einschätzung entsteht. Umso überraschender, dass 116 TN tendenziell einer a priori erfolgten Zuschreibung von Qualität zu Dokumenten/ Items positiv gegenüber stünden.

Akzeptanz von Empfehlungsdiensten

Die Akzeptanz von Empfehlungsdienste sinkt und steigt mit deren Qualität. Bei einer 90%-Treffergenauigkeit des Empfehlungsdienstes wird dieser von 192 TN (75,9 %) als positiv (häufig bzw. immer nützlich) bewertet, bei einer Genauigkeit von 50% von 41 TN (20,1%).

Wie bewerten Sie den Nutzen eines Empfehlungsdienstes, der in neun von zehn Fällen einen guten Vorschlag macht?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	6	2,4	2,4	2,4
selten nützlich	4	1,6	1,6	4,1
manchmal nützlich	43	17,0	17,6	21,6
häufig nützlich	108	42,7	44,1	65,7
immer nützlich	84	33,2	34,3	100,0
Gesamt	245	96,8	100,0	
Fehlend	8	3,2		
Gesamt	253	100,0		

Wie bewerten Sie den Nutzen eines Empfehlungsdienstes, der in fünf von zehn Fällen einen guten Vorschlag macht?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	18	7,1	7,3	7,3
selten nützlich	67	26,5	27,3	34,7
manchmal nützlich	109	43,1	44,5	79,2
häufig nützlich	39	15,4	15,9	95,1
immer nützlich	12	4,7	4,9	100,0
Gesamt	245	96,8	100,0	
Fehlend	8	3,2		
Gesamt	253	100,0		

Rankingoptionen

Die Verwendung von quantitativen Nutzungsinformationen (Downloadhäufigkeiten) zum Dokumentranking wird tendenziell als wenig gewinnbringend erachtet. 34,8% der TN erachteten solche Funktionen als gar nicht oder selten sinnvoll, 25,3% als häufig oder immer sinnvoll.

Für wie nützlich halten Sie eine Sortierung der Trefferliste anhand der Nutzungszahlen der Dokumente?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	26	10,3	10,5	10,5
selten nützlich	62	24,5	25,1	35,6
manchmal nützlich	95	37,5	38,5	74,1
häufig nützlich	44	17,4	17,8	91,9
immer nützlich	20	7,9	8,1	100,0
Gesamt	247	97,6	100,0	
Fehlend	6	2,4		
Gesamt	253	100,0		

Höhere Akzeptanz finden Rankingoptionen, die auf Zitationshäufigkeiten basieren: Nur 12,7 % stuften diese als selten oder gar nicht nützlich ein, wohingegen 51 % sie als häufig oder immer nützlich betrachteten.

Für wie nützlich halten Sie eine Sortierung der Trefferliste danach wie oft die Dokumente von anderen zitiert wurden?

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	6	2,4	2,4	2,4
selten nützlich	26	10,3	10,5	13,0
manchmal nützlich	86	34,0	34,8	47,8
häufig nützlich	86	34,0	34,8	82,6
immer nützlich	43	17,0	17,4	100,0
Gesamt	247	97,6	100,0	
Fehlend	6	2,4		
Gesamt	253	100,0		

Die Anzeige der Gesamtnutzung eines Repositories scheint für Wissenschaftler etwas weniger interessant zu sein, als sie es für TN aus dem Bibliotheksumfeld ist. Während 32,4% aller TNs diese Information als häufig oder immer nützlich erachteten, traf dies nur auf 23% der teilnehmenden Wissenschaftler zu.

Bewerten Sie den Nutzen einer Anzeige der Gesamtnutzung eines Repositoriums.

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig				
gar nicht nützlich	17	6,7	7,0	7,0
selten nützlich	60	23,7	24,8	31,8
manchmal nützlich	83	32,8	34,3	66,1
häufig nützlich	49	19,4	20,2	86,4
immer nützlich	33	13,0	13,6	100,0
Gesamt	242	95,7	100,0	
Fehlend	11	4,3		
Gesamt	253	100,0		

Bewerten Sie den Nutzen einer Anzeige der Gesamtnutzung eines Repositoriums [Im Bibliothekswesen Tätige ausgeschlossen]

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	gar nicht nützlich	11	8,1	8,5	8,5
	selten nützlich	39	28,9	30,2	38,8
	manchmal nützlich	48	35,6	37,2	76,0
	häufig nützlich	19	14,1	14,7	90,7
	immer nützlich	12	8,9	9,3	100,0
	Gesamt	129	95,6	100,0	
Fehlend		6	4,4		
Gesamt		135	100,0		

Zusammenfassung

Aufgrund der beruflichen Tätigkeit der Teilnehmer der Onlinebefragung (118 TN, die im Bibliothekswesen tätig waren, bei 253 TN insgesamt) ist von einer Verzerrung der Ergebnisse auszugehen. Ausgehend von Ergebnissen der Experteninterviews wurden die Repositorynutzer nach möglichen Funktionalitäten in Repositorien befragt.

Positiv bewertet wurden v.a. Recommenderdienste, die Empfehlungen formulieren

- auf Basis des eigenen Lese-/ Nutzungsverhaltens (46,7% der TN bewerteten diese als häufig oder immer nützlich)
- auf Basis der an den Publikationen beteiligten Autoren (39,5% der TN bewerteten diese als häufig oder immer nützlich)
- auf Basis eines Abgleichs des eigenen Lese-/ Nutzungsverhaltens mit dem Lese-/ Nutzungsverhaltens anderer Nutzer (36,4% der TN bewerteten diese als häufig oder immer nützlich)

Übereinstimmend mit den Experten verlangen auch die Nutzer eine hohe Qualität der Recommender: Während 75,9% der antwortenden TN Empfehlungsdienste mit einer 90%igen Treffergenauigkeit als häufig oder immer positiv bewerten⁴, bewerteten nur 20,1% der antwortenden TN einen Service als häufig oder immer nützlich, der mit 50%iger Präzision passende Dokumente empfiehlt⁵.

Den Informationsgehalt von Zugriffs- und Downloadstatistiken bewerteten 47,8% der TN als häufig oder immer nützlich. Die Verwendung rein quantitativer Daten (absolute Nutzungshäufigkeiten) zur Abschätzung der Qualität eines Dokumentes wurde allerdings sehr negativ gesehen (52,2 % der antwortenden TN bewerteten diese Option als gar nicht oder selten nützlich, nur 15,5 % bewerteten sie als häufig oder immer nützlich). Interessanterweise gibt es also eine große Divergenz zwischen dem wahrgenommenen Wert von Nutzungshäufigkeiten und dem Glauben daran, aus der puren Nutzungshäufigkeit von Dokumenten ihre Qualität ableiten zu können. Es lässt sich vermuten, dass die quantitative Nutzungsinformation in der Wahrnehmung der Nutzer relevante Aussagen (47,8% erachten sie als überwiegend nützlich) über eine andere Dimension als *Qualität* trifft. Dies

⁴ Immerhin 34,3% der antwortenden TN bewerteten einen solchen Service sogar als immer nützlich.

⁵ Ein Service mit 50%iger Genauigkeit wurde nur von 4,9% der antwortenden TN als immer nützlich bezeichnet.

korrespondiert mit den Ergebnissen der Experteninterviews, in denen einer der beiden Wissenschaftler absolute Nutzungshäufigkeiten als wichtiges Feedbackinstrument betrachtete, allerdings überwiegend vor einer Überinterpretation quantitativer Daten gewarnt wurde.

Auch die Ergebnisse auf die Frage, wie nützlich die Nutzer die Sortierung der Trefferlisten nach Nutzungszahlen der Dokumente halten, widersprechen dem wahrgenommenen Nutzen von Downloadstatistiken: Nur 25,3 % der TN bewerteten solche Sortierungsfunktionen als häufig oder immer nützlich. Eine mögliche Sortierung nach Zitationshäufigkeiten stieß auf größere Zustimmung, 51% der TN bewerteten sie als häufig oder immer nützlich. Etwas höher bewertet wurde die Anzeige der Gesamtnutzung des Repositories: 32,4 % der TN bewerteten diese als häufig oder immer nützlich, allerdings bewerteten 30,4% sie als selten oder gar nicht nützlich. Dieses Ergebnis steht etwas im Widerspruch zur Wertschätzung, die diese Information von den Repositorybetreibern erfährt. Allerdings haben diese auch spezifische Verwendungszecke für die Informationen (Monitoring, Optimierung). Eine geringere Übereinstimmung zwischen Expertenaussagen und Nutzerbewertungen gab es hinsichtlich der Verwendung von Kontextinformationen der Nutzung. Diese wurden von mehreren Experten als wertvoll für die Verbesserung der Services angesehen. Auch wenn die Bewertung verschiedener Kontextinformationen durch die Nutzer besser ausfällt, als es bei der Bewertung der absoluten Nutzungshäufigkeiten der Fall ist, divergiert die Einschätzung von der Expertenmeinung:

- 21,3% der TN bewerteten die Anzeige der Information, wie häufig ein Dokument gebookmarkt wurde, als häufig oder immer nützlich
- 15,1% der antwortenden TN bewertenden die Messung der Verweildauer der Nutzer auf einem Dokument als häufig oder immer nützlich

Vermittelnd muss angemerkt werden, dass vermutlich das Nutzerurteil über Services, die diese Informationen auswerten, entscheidender sein dürfte als eine Einschätzung der von den Services ausgewerteten Informationen. Das von den meisten Nutzern wohl geschätzte Google-Ranking wertet etwa u.a. die Verweildauer auf Dokumenten mit aus.⁶

Positiver bewertet wurden explizit qualitative Statements über Dokumente, z.B.

- Nutzerbewertungen in Form einer Notenskala: 29,3% der TN bewerteten diese als häufig oder immer nützlich
- Nutzerbewertungen durch kurze öffentliche Kommentare: 46,7% der TN bewerteten diese Funktion als häufig oder immer nützlich

Dieses Ergebnis deckt sich damit sehr stark mit der Einschätzung von SeExp (Suchmaschinenexperte). Generell scheint bei Nutzern ein großes Interesse an Hinweisen auf die Qualität von Dokumenten vorhanden zu sein. 45,9% der TN bewerten solche Funktionen positiv (häufig oder immer nützlich), nur 14,7% bewerteten sie als selten oder gar nicht nützlich.⁷

⁶ Auch wenn dieser nur zur Illustration genannte Faktor für wissenschaftliche Dokumente schlecht herangezogen werden kann, da diese meist als PDF-Datei vorliegen und daher oft offline oder ausgedruckt genutzt werden dürften.

⁷ Da diese Frage ohne nähere Beschreibung oder Operationalisierung des Konstrukts *Qualität* gestellt wurde, überrascht die positive Bewertung.

4. Befragung der Repositorybetreiber

Im September 2009 versandte die SULB den Fragebogen an die Betreiber aller Repositories, die zu diesem Zeitpunkt in der Repositoryliste der Deutschen Initiative für Netzwerkinformation DINI e.V. gelistet waren.⁸ Insgesamt wurden Mails an die Betreiber von 130 Repositories verschickt. Es gingen Antworten aus 17 Einrichtungen ein, wovon eine Einrichtung die Teilnahme an der Befragung ablehnte. Die Antworten der verbleibenden 16 Einrichtungen gaben Auskünfte über 26 Repositories⁹ und auf 16 von einer Einrichtung gehostete Journals. Auch wenn Journalplattformen nicht im Fokus der Umfrage standen, werden die Antworten in den Auswertungen berücksichtigt. Letztlich standen für die Auswertung Informationen über 42 Publikationsservices (davon 16 Journale, 26 Repositories) aus 16 Einrichtungen zur Verfügung.

Ergebnisse

Erhebung und Anzeige von Benutzerzugriffen auf den Publikationsservices

Die in Frage 1 des Bogens erhobenen Variablen 1-3 sollten in Erfahrung bringen, ob auf dem Server des Publikationsangebotes Informationen über Benutzerzugriffe im Sinne einer serverweiten Zugriffsstatistik erhoben werden und für welche Nutzergruppen (interne Nutzer/ Admins oder externe Nutzer/ Besucher der Seite) diese zugänglich ist.

Insgesamt 39 der 42 Publikationsservices (Repositories plus Journals) erfassten, auf welche Informationen die Benutzer zugriffen.

Erhebung der Benutzerzugriffe

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	3	7,1	7,1	7,1
	Ja	39	92,9	92,9	100,0
	Gesamt	42	100,0	100,0	

Bei 13 der Services wurden die Informationen zur internen Anzeige genutzt, 23 verzichteten darauf, zu sechs der Services gab es keine weiteren Informationen.

Interne Anzeige der Benutzerzugriffe

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	23	54,8	63,9	63,9
	Ja	13	31,0	36,1	100,0
	Gesamt	36	85,7	100,0	
Fehlend	K.A.	6	14,3		
Gesamt		42	100,0		

⁸ Online unter: [http://www.dini.de/wiss-publizieren/repository/?tx_repository_pi1\[mode\]=name](http://www.dini.de/wiss-publizieren/repository/?tx_repository_pi1[mode]=name)

⁹ Teils hosteten Einrichtungen oder betrieben mehrere Repositories

Extern (für Nutzer) zeigten elf der Services die Informationen über Nutzergriffe an, 31 Services nutzten diese Möglichkeit nicht.

Externe Anzeige der Benutzerzugriffe

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	31	73,8	73,8	73,8
	Ja	11	26,2	26,2	100,0
	Gesamt	42	100,0	100,0	

Erhebung und Anzeige von Benutzerverhalten auf den Publikationsservices

Frage 2 (mit den Variablen vier bis sechs) zielte darauf, ob auf einem Publikationsservices Daten über das Verhalten von Nutzern im Sinn einer Clickstreamanalyse erhoben werden und ob eine Anzeige dieser Daten erfolgt.

Nach Betreiberantwort erfolgt auf 25 der Publikationsservices eine solche Erhebung des Benutzerverhaltens. Da aber nur bei acht der Services eine interne und bei einem Service eine externe Anzeige der Daten erfolgt, ist zu vermuten, dass sehr selten eine Analyse der Daten vorgenommen wird und die 25 bestätigenden Antworten auf Frage 4 sich auf die unvermeidliche Sammlung der Daten in Webserverlogs beziehen.

Erhebung von Benutzerverhalten

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	17	40,5	40,5	40,5
	Ja	25	59,5	59,5	100,0
	Gesamt	42	100,0	100,0	

Interne Anzeige des Benutzerverhaltens

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	34	81,0	81,0	81,0
	Ja	8	19,0	19,0	100,0
	Gesamt	42	100,0	100,0	

Externe Datenanzeige des Benutzerverhaltens

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	41	97,6	97,6	97,6
	Ja	1	2,4	2,4	100,0
	Gesamt	42	100,0	100,0	

Erhebung und Anzeige von Dokumentzugriffen auf den Publikationsservices

Frage 3 (Variablen sechs bis neun) befasste sich mit der Messung von Nutzerzugriffen auf Dokumente und damit einem der wesentlichen Themen des Open-Access-Statistik-Projekts:

Erhebung der Zugriffe auf einzelne Dokumente

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	4	9,5	11,1	11,1
	Ja	32	76,2	88,9	100,0
	Gesamt	36	85,7	100,0	
Fehlend	K.A.	6	14,3		
Gesamt		42	100,0		

Interne Anzeige der Dokumentzugriffe

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	11	26,2	30,6	30,6
	Ja	25	59,5	69,4	100,0
	Gesamt	36	85,7	100,0	
Fehlend	K.A.	6	14,3		
Gesamt		42	100,0		

Externe Anzeige der Dokumentzugriffe

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	16	38,1	38,1	38,1
	Ja	26	61,9	61,9	100,0
	Gesamt	42	100,0	100,0	

32 der 42 Services gaben an Nutzerzugriffe auf einzelne Dokumente zu erheben, vier verneinten dies, sechs machten keine Angaben. 25 Services nutzen diese Informationen zur internen Anzeige, 26 zur externen Anzeige.

Technische Lösungen, Standardisierung und Behandlung von Verzerrungen

Mit den Fragen 4 bis 7 sollten Informationen darüber gesammelt werden, mit welchen Softwarelösungen die Nutzungsdaten ausgewertet werden und ob dabei ein Standard befolgt wird. In Fällen, in denen kein Standard angewandt wird, wurde erfragt ob der Betreiber des Services sich bemüht, Verzerrungen wie etwa maschinelle Zugriffe (durch indexierende Suchmaschinen oder Webharvester) oder Doppelclicks (Aufruf desselben Dokuments innerhalb kurzer Zeit) zu kontrollieren oder zu eliminieren.

Frage 4: Verwendete Software

	AWSTATS	Analog	DSpace	EPrints	OPUS	Eigenes Skript	Summe
Services	27	1	1	1	4	5	39
Betreiber	6	1	1	1	1	5	15

Am häufigsten kommt die Software AWSTATS zur Erstellung von Nutzungsstatistiken zum Einsatz. Die Software liefert AWSTATS in der Standardkonfiguration serverweite Nutzungsstatistiken, kann nach entsprechender Konfiguration aber auch dokumentbezogene Statistiken anbieten. Einmal fand die Software Analog Anwendung, sechs der Repositories verwenden die Funktionen, die ihre Softwareumgebung (OPUS, DSpace, EPrints) bereitstellt und fünf vertrauen selbstentwickelten Programmskripten. Allerdings fallen alleine 22 der AWSTATS-Anwendungen auf Hostinganbieter, dasselbe gilt für alle Nennungen der OPUS-internen Statistikfunktion.

Frage 5: Anwendung eines Standards

		Häufigkeit	Prozent	Gültige Prozente
Gültig	Nein	34	81,0	87,2
	Ja	5	11,9	12,8
	Gesamt	39	92,9	100,0
Fehlend	K.A.	3	7,1	
Gesamt		42	100,0	

Fünf der Services gaben an, Nutzungsdatenerhebung nach den Vorgaben eines Standards zu erheben, in allen dieser Fälle handelte es sich um dabei um die COUNTER Codes of Practice¹⁰.

Frage 6 sollte herauszufinden, ob Services, die keinen Standard anwenden, sich zumindest um eine Filterung der erwähnten maschinellen Zugriffe bemühen.

Frage 6: Filterung maschineller Zugriffe, wenn kein Standard angewandt wird

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	7	16,7	20,0	20,0
	Ja	28	66,7	80,0	100,0
	Gesamt	35	83,3	100,0	
Fehlend	K.A.	7	16,7		
Gesamt		42	100,0		

¹⁰ http://www.projectcounter.org/code_practice.html

Für immerhin 28 der 35 Repositorien¹¹, zu denen eine Angabe verfolgte, wurde zumindest eine Filterung (auch wenn sie nicht nach einem Standard erfolgt) bejaht.

Frage 7 richtete sich analog darauf, ob (in Fällen, in denen kein Standard zur Anwendung kommt) zumindest Versuche unternommen werden Doppelclicks auf Dokumente zu kontrollieren.

Frage 7: Ermittlung von Doppelclicks

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Nein	9	21,4	25,7	25,7
	Ja	26	61,9	74,3	100,0
	Gesamt	35	83,3	100,0	
Fehlend	K.A.	7	16,7		
Gesamt		42	100,0		

Von den 35 Services, zu denen Angaben erfolgten, wurde die Eliminierung von Doppelclicks in 26 Fällen bejaht. In Fällen, in denen Doppelclicks nicht nach Vorgaben eines Standards eliminiert werden, wurde nach der Dauer der Doppelclickspanne¹² gefragt: Überraschenderweise gingen nur zu drei der 26 Services, die in dieser Art Doppelclicks eliminieren, Informationen zur Doppelclickspanne ein (unmittelbar aufeinanderfolgende Zugriffe, 60 Sekunden sowie 30 Minuten).

Die restlichen drei Fragen des Bogens waren offen formuliert und richteten sich

- auf Mindestanforderungen, die seitens des Servicebetreibers an Nutzungsstatistiken gestellt werden (Frage 8)
- auf Effekte, die sich der Servicebetreiber von einer Nutzungsstatistik erhofft (Frage 9) und
- auf nutzungsdatenbasierten Zusatzfunktionalitäten, die der Servicebetreiber wünscht (Frage 10).

Da diese Fragen auf Einschätzungen eines Servicebetreibers zielen und nicht auf die Ausstattung von Publikationsservices, verringert sich die Anzahl der Untersuchungsobjekte von 42 auf 16. Bei allen Fragen waren Mehrfachauskünfte pro Betreiber möglich.

¹¹ Die Zahl von 35 Services, zu denen eine Angabe auf die Frage nach der Elimination von Doppelclicks ohne Verwendung eines Standards erfolgte, und der Zahl 34 für Services, die keinen Standard anwenden, erklärt sich daraus, dass auf Frage 7 auch Antworten eines Services eingingen, der zu Frage 5 keine Auskünfte gab.

¹² Die Zeit, die zwischen zwei Aufrufen eines Dokuments liegen muss, damit beide Zugriffe in die Auswertung miteingehen.

Frage 8: Mindestanforderungen der Servicebetreiber an Nutzungsstatistiken

Anforderung	Nutzungsmessung (Dokumente Ebene)	Eliminierung von Verzerrungen	Granularität der Daten	Serverweite Statistiken
	13	5	2	2

Die am häufigsten genannte Anforderung an Nutzungsstatistiken war die exakte Messung der Dokumentnutzung (13 Antworten), gefolgt von der Forderung nach Eliminierung von Verzerrungen (Spider, maschinelle Downloads, Doppelclicks) mit 5 Antworten. Je zweimal genannt wurden die Granularität der Daten sowie die Forderung nach zusätzlichen serverweiten Statistiken. Je einmal genannt wurden:

- Manipulationssicherheit
- Impact Messung
- Beachtung datenschutzrechtlicher Regelungen
- Auswertungen nach verschiedenen Standards
- Ausweisen nicht nur der Volltextnutzung, sondern auch der Nutzung von Abstractpages (Frontdoors, Splashpages), die den Volltexten meist vorgeschaltet sind
- Statistik auf Dokumente Ebene, die Zugriffe auf einzelne Teile von Mehrdateiendokumenten summiert
- Trennung einer internen Adminsicht von einer externen Sicht für Benutzer
- Auskunft über Ansicht der fehlgeschlagenen Dokumentaufrufe

Frage 9: Von einer Nutzungsstatistik erhoffte Effekte

Effekte	Werbung	Performance- messung	Autorenservice	Argumentations- hilfe pro OA	Attraktivität	Nutzungsanalyse	Vergleichbarkeit von Repositories
	8	7	5	5	4	3	3

Acht der Serviceprovider versprochen sich einen Werbeeffect von der Einführung einer Nutzungsstatistik, sieben eine Performancemessung (oder Reichweitenmessung) des Publikationsservices und der darauf zugänglichen Dokumente, fünf erachteten die Statistik als nützlichen Autorenservice, vier versprochen sich von ihr eine Steigerung der Attraktivität des Services, vier erachteten sie als Werkzeug zur Nutzungsanalyse und drei erhofften sich eine

Vergleichbarkeit von Repositories. Je einmal wurden als von einer Nutzungsstatistik erhoffte Effekte genannt:

- Aufschluss über Akzeptanz verschiedener Dokumentarten
- Universitätsinterne Argumentationshilfe für Repositorybetrieb (Evaluation)

Frage 10: Erwünschte nutzungsdatenbasierte Zusatzfunktionalitäten

Funktionalitäten	Kontextinformationen	Dokumentranking	VG WORT Unterstützung	Zitationsstatistik	Freie Gestaltbarkeit der Anzeige	Recommender
	9	6	3	2	2	2

Am häufigsten nannten die Servicebetreiber die Gewinnung von Kontextinformationen der Dokument- und Repositorynutzung als wünschenswertes Zusatzfeature. Unter Kontextinformationen wurden folgende Angaben subsummiert:

- Aufschlüsselung der nationalen Herkunft der Zugriffe
- Zeitliche Differenzierung nach Tageszeiten, Wochentagen, Monaten
- Differenzierung nach Browsern und Betriebssystemen
- Identifikation des Nutzungskontext (Suchmaschine, verwendete Suchbegriffe)
- Identifikation von Ein- und Ausstiegsseiten
- Analyse der Nutzungswege (Clickstreams)

Die am zweithäufigsten gewünschte Zusatzfunktion ist ein nutzungsdatenbasiertes Dokumentranking, gefolgt von einer Unterstützung des VG Wort Mechanismus zur Bestimmung der Zugriffe auf elektronische Dokumente. Zwei Betreiber wünschten parallele Zitationsanalysen und ebenfalls zwei Antworten forderten Recommender-/Empfehlungsdienste sowie die freie Gestaltbarkeit der Anzeige der Nutzungsstatistiken.

Je einmal genannt wurden:

- Vergleichbarkeit durch Standardisierung
- Statistiken auf Artekelebene
- Statistiken auf Autorenebene
- Einblendung der meistgenutzten Dokumente
- Nutzungsmetriken
- Auswertung von Linkresolver-Logfiles
- Berücksichtigung nur vollständiger Downloads
- Aufzeichnung der Nutzungsdauer eines Dokuments
- Reportfunktion, die Dokumentnutzung nach Dokumenttyp und Institution (Fachbereich, Lehrstuhl) ausweist

Zusammenfassung

Die antwortenden Betreiber von Repositorien resp. Journalen schätzen den Wert von Nutzungsinformationen offensichtlich hoch ein: Auf 39 der 42 Plattformen werden serverweite, nicht dezidiert dokumentbezogene Benutzerzugriffe erhoben. Bei 13 Services werden die Zugriffsinformationen intern angezeigt, offensichtlich zum institutionseigenen Monitoring, elf Services zeigen die Informationen auch für externe Nutzer an. Ein größerer Anteil der Services beschränkt sich anscheinend darauf Daten vorzuhalten, stellt sie aber weder intern noch extern dar.

Ähnlich verhält es sich bei Frage nach der Erfassung des Benutzerverhaltens in Form von Clickstream-Analysen. Für 25 Services wurde angegeben, dass eine solche Erfassung erfolgt. Aufgrund der Angaben zur internen (8) oder externen (1) Anzeige der Informationen lässt sich allerdings vermuten, dass die Informationen hierzu i.d.R. in Form von Webserverlogs und daher nicht aufbereitet vorliegen.

An einem der Kernpunkte des Projekts Open-Access-Statistik herrscht offensichtlich großes Interesse: Die Frage nach der Erfassung von dokumentbezogenen Zugriffsinformationen wurde für 32 der Services bejaht, bei 25 Services wurden diese Informationen intern angezeigt, bei 26 extern. Da eines der wichtigsten Anliegen von OAS Standardisierung und Vergleichbarkeit der Nutzungsdaten sind, folgte ein Fragenkomplex zu Technik und Standards. Bei der Software, die verwendet wurde herrscht eine eindeutige Dominanz der Software AWStats, die allerdings genauso wenig wie die anderen genannten Lösungen einen Standard darstellt oder eine Vergleichbarkeit der Daten garantiert. Dies spiegeln auch die Ergebnisse der Frage nach den verwendeten Standards wider: Nur fünf der Services wenden bei der Datenaufbereitung nach Betreiberauskunft einen Standard an, in allen dieser Fälle handelt es sich um den Standard COUNTER. Betreiber, die keinen Standard zur Aufbereitung der dokumentbezogenen Nutzungsdaten anwenden, unternehmen dennoch zum Großteil (28 Services) Versuche, Zugriffe durch Leser von maschinellen Zugriffen (wie sie etwa bei der Indexierung durch Suchmaschinenroboter entstehen) zu unterscheiden und die Anzeige der Dokumentzugriffe dahingehend zu bereinigen. Ähnlich verhält es sich mit einer Behandlung einer anderen Art von Verzerrung, die Dokumentzugriffsstatistiken betreffen: Für 26 der Repositories wurde angegeben, dass bei der Messung der Dokumentzugriffe Doppelclicks¹³ eliminiert werden – allerdings ohne Verwendung eines Standards und daher nicht in vergleichbarer Art.

Im folgenden Block mit offenen Fragen konnten die Plattformbetreiber Anforderungen an Nutzungsstatistiken formulieren, welche Effekte sie sich von der Einführung einer Nutzungsstatistik erhoffen und äußern, welche nutzungsdatenbasierten Zusatzfunktionen sie sich für ihre Plattform wünschen. 13 der 16 Servicebetreiber legten Wert auf dokumentbezogene Nutzungsstatistiken, oft (fünffmal) inkl. der Bereinigung von Verzerrungen (maschinelle Zugriffe, Doppelclicks). In den Kontext der zu eliminierenden Verzerrungen könnte man auch den an anderer Stelle geäußerten Punkt der ausschließlichen Berücksichtigung vollständiger Downloads nennen.

Bei den erhofften Effekten standen mit Werbung (8 Nennungen), Performancemessung/ Reichweitenmessung (7) und Schaffung einer Argumentationshilfe pro Open Access (5) Wirkungen im Vordergrund, die v.a. der eigenen Organisation bzw. dem eigenen Service dienen. Überraschend oft wurden Effekte genannt, die Monitoringcharakter haben: Neben der erwähnten

¹³ Also das Aufrufen desselben Dokuments in kurzer zeitlicher Abfolge durch denselben Nutzer. In der Regel werden solche Doppelclicks innerhalb einer definierten Zeitspanne eliminiert, da man davon ausgeht, dass innerhalb dieses sog. Doppelclickintervalls kein erneuter Nutzungsvorgang vorliegen kann.

Performancemessung, Nutzungsanalyse (3) und die Herstellung von Vergleichbarkeit von Repositories (3). Zu diesem Sammelpunkt könnte man auch die einmal genannte Schaffung einer inneruniversitären Argumentationshilfe oder den an anderer Stelle genannten Wunsch nach Informationen über fehlgeschlagene Dokumentzugriffe nennen. Eher nach außen gerichtete Effekte wurden seltener genannt, etwa die Schaffung eines Autorenservices (3).

Die Liste der erwünschten nutzungsdatenbasierten Zusatzfunktionen führen ähnlich Dienste an, die Kontextinformationen (9) der Nutzung berücksichtigen. Die Liste der zu berücksichtigenden Kontextinformationen lässt dabei auf sehr unterschiedliche Verwendungsabsichten schließen, im Vordergrund scheint aber neben der Entwicklung von Nutzungsnetzwerken v.a. die Nutzungsanalyse der Repositories zur Optimierung des Services zu stehen. So wünschen sich die Betreiber v.a. Informationen über nationale Herkunft der Services, Zeitschema der Nutzung nach Wochentagen/Uhrzeiten, Informationen über Browser/Betriebssysteme, Informationen über verwendete Suchmaschinen und –begriffe, Identifikation der Ein-/Ausstiegsseiten oder die Aufzeichnung der Nutzungsdauer einzelner Dokumente.

Am zweithäufigsten nach der Verarbeitung der Kontextinformationen wünschten sich die Betreiber Dokumentrankings (6 Nennungen), die Unterstützung der VG Wort Technik zur Messung der Dokumentzugriffe (3), Zitationsstatistiken (2), eine freie Gestaltbarkeit der Statistikanzeige und Recommendersysteme (je 2).

Bei den offenen Fragen (8-10) wurden vereinzelt auch Aspekte thematisiert, die sich nicht sauber einem der Bereiche Anforderungen/ Effekte/ Zusatzfunktionalitäten zuordnen lassen. Dazu zählen:

- Manipulationssicherheit der Nutzungsdaten
- Auswertung der Nutzungsinformationen nach verschiedenen Standards
- Nutzungsmetriken und Impact Messung durch Nutzungsdaten
- Datenschutzkonforme Techniken und Auswertungen
- Messung der Volltext- und Abstractpage-Zugriffe
- Standardisierung
- Nutzungsstatistiken auf Artikelebene
- Nutzungsstatistiken auf Autorenebene
- Reportfunktionen, die Nutzungszahlen nach Dokumenttypen und organisatorischer Einheiten (Institut, Fachbereich, Lehrstühle etc.) nachweisen
- Einblendung der meistgenutzten Dokumente
- Auswertung von Nutzungsinformationen aus Linkresolvern

Bilanz

Dem hohen Bedarf an der Erhebung von (dokumentbezogenen oder serverweiten) Nutzungsstatistiken steht ein Mangel an Standardisierung gegenüber. Dies trifft sowohl auf die Standardisierung der Parameter mittels derer Zugriffszahlen ermittelt werden, als auch auf die Behandlung der Verzerrungen (Definition von Doppelclickintervallen, Vorgehen bei der Eliminierung maschineller Zugriffe) zu. Insbesondere fehlt es an Interoperabilität und Vergleichbarkeit der Statistiken oder gar an Mechanismen zur serverübergreifenden Sammlung von Nutzungsinformationen. Die bevorzugten Szenarien zur Verwendung der Nutzungsdaten entsprechen klar der Perspektive von Infrastruktur- und Serviceanbietern: Im Vordergrund stehen Verwendungen zum Monitoring des Angebots und zu dessen Optimierung. Explizit an externe Partner adressiert sind

Verwendungsszenarien wie Argumentationsunterstützung pro Open Access/ Werbung (Adressat: Nutzer), Services für Autoren, Rechtfertigungshilfe für den Betrieb der Plattform (Adressat: Universitätsverwaltung). Die Zusatzfunktionen sollen v.a. Mehrwerte durch Berücksichtigung der Kontextinformationen liefern, scheinen aber v.a. den erwähnten Monitoringzwecken unterstellt zu sein und der Optimierung der Dienstleistung zu dienen. Dienstleistungen, die für Nutzer oder Autoren relevant sind, werden eher nachrangig erwähnt (Dokumentrankings, Recommender).

5. Fazit

Die drei Teilnehmergruppen (Experten, Nutzer, Servicebetreiber) nannten teils unterschiedliche Funktionen oder Informationen, die Publikationsservices ihrer Meinung nach anbieten sollten. Zusätzlich wurden manche der Funktionen und Informationen von den drei Gruppen auch unterschiedlich bewertet. Auch wenn im Mittelpunkt der Erhebungen Funktionen standen, die auf Nutzungsdaten¹⁴ beruhen, wurden auch Einschätzungen zu anderen Funktionen diskutiert.

In dieser Zusammenfassung wird daher nach primär nutzungsdatenbasierten Anwendungen und nicht-nutzungsdatenbasierten Anwendungen unterschieden. Verwandte Anwendungen werden in Blöcken zusammengefasst und ihnen werden die jeweiligen Ergebnisse der Erhebungen zugeordnet, Details finden sich in den vorangehenden Ausführungen zu den einzelnen Erhebungen.

In den Experteninterviews wurden ein Suchmaschinenexperte, ein Retrievalexperte, ein Repositoryexperte, eine Wissenschaftlerin und ein Wissenschaftler befragt. An der Onlinebefragung nahmen insgesamt 253 Personen teil, von denen zum Zeitpunkt der Befragung 118 im Bibliothekswesen tätig waren.

Verwendung absoluter Nutzungshäufigkeiten von Dokumenten

Die interviewten *Experten* bewerten die Anzeige von Nutzungshäufigkeiten einzelner Dokumente positiv. Sie betonen allerdings auch, dass der Fachkontext der Dokumente und Wissenschaftler sowie andere Variablen berücksichtigt werden müssen. Drei der Experten schlugen vor, nicht nur die absolute Nutzungshäufigkeit von Dokumenten anzuzeigen, sondern auch die Abstractviews zu zählen. Sie gingen davon aus, vielleicht aus dem Verhältnis von Abstractview und Volltextdownload Aufschluss über die Qualität des Volltexts zu erlangen.¹⁵

Auch die *Teilnehmer der Onlinebefragung* bewerteten den Informationsgehalt von Zugriffs- und Downloadstatistiken hoch (47,8% als häufig oder immer nützlich), ihre Anzeige bei einzelnen Dokumenten bewerteten 37,1% als häufig oder immer nützlich. Als Indikator für Qualität erachteten die Teilnehmer der Onlinebefragung die absolute Nutzungshäufigkeiten einzelner Dokumente eher nicht: 52,2 % der antwortenden TN bewerteten diese Option als gar nicht oder selten nützlich. Aus dieser Divergenz der hohen Wertschätzung der Informationen über die quantitative Dokumentnutzung und dem Misstrauen, daraus Qualität der Dokumente ableiten zu können, lässt sich die Vermutung ableiten, dass die quantitative Nutzungsinformation hohe Relevanz für eine andere Dimension als *Qualität* besitzt. Dementsprechend bewerten auch nur 25,3 % der Teilnehmer solche Sortierungsoptionen nach Downloadhäufigkeiten als häufig oder immer nützlich. Zwiespältig bewertet wurde die Anzeige der Gesamtnutzung des Repositories: 32,4 % erachteten sie als häufig oder immer nützlich.¹⁶ Allerdings bewerteten auch 30,4% sie als selten oder gar nicht nützlich. Dieses Ergebnis steht etwas im Widerspruch zur höheren Wertschätzung, die diese Information von den *Servicebetreibern* erfährt. Allerdings haben diese auch spezifische Verwendungszecke für die Informationen (Monitoring, Optimierung).

¹⁴ Genauer: Informationen über die Nutzung wissenschaftlicher Onlinedokumente

¹⁵ Ein Experte merkte einschränkend an, dass ein niedriges Verhältnis von Volltextdownloads zu Abstractview ggf. eher auf einen uninteressanten Abstract schließen lassen könnte.

¹⁶ Schließt man die im Bibliothekswesen Tätigen aus, fanden nur noch 23% der TN diese Anzeige häufig oder immer nützlich.

Die antwortenden Betreiber von *Publikationsservices* (Repositorien, Journalen) haben offensichtlich großes Interesse an der Messung der Zugriffe: 39 der 42 Plattformen erheben serverweite Benutzerzugriffe. Auch eines der wesentlichen OAS-Themen hat hohe Relevanz für die Servicebetreiber: 32 der Services erhoben dokumentbezogene Zugriffsstatistiken, 25 zur internen, 26 zur externen Anzeige. Die Betreiber der Services äußerten auch den Wunsch nach dokumentbezogenen Nutzungsstatistiken inkl. der Bereinigung von Verzerrungen. Sechs der 16 Betreiber wünschten sich Rankingoptionen nach Nutzungshäufigkeiten.

Standardisierung

Der recht weitverbreiteten Verwendung von Nutzungshäufigkeiten steht allerdings ein deutlicher Mangel an Standardisierung und Interoperabilität gegenüber: Nur bei fünf der Publikationsservices wurde ein Standard (in allen Fällen COUNTER) zur Ermittlung der Zugriffshöhen angewandt. 28 der Betreiber, die keinen Standard verwenden, versuchen mit eigenen Mitteln maschinelle Zugriffe aus den Statistiken zu filtern. 26 dieser Betreiber versuchen Doppelclicks, Verzerrungen durch rasches Mehrfachaufrufen von Dokumenten innerhalb einer Sitzung, zu eliminieren. Den fünf Services, die einen Standard zur Ermittlung von Zugriffszahlen nutzen, stehen also neun Services, die keine maschinellen Zugriffe eliminieren, und elf Services, die keine Doppelclicks eliminieren, entgegen sowie 28 resp. 26 Services, die solche Verzerrungen unstandardisiert eliminieren. Folglich werden keine vergleichbaren Daten generiert, die Basis weitergehender Auswertungen werden können.

Das Bedürfnis an solchen homogenen Daten ist bei den Servicebetreiber allerdings nicht zu leugnen, sie nannten als Anforderungen an dokumentbezogene Nutzungsdaten u.a. folgende korrespondierenden Eigenschaften: Manipulationssicherheit der Nutzungsdaten, Auswertung nach verschiedenen Standards, Metriken, Datenschutzkonformität.

Einbeziehung von Kontextinformationen und Bewertungen

Die Verwertbarkeit und Aussagekraft reiner Nutzungshäufigkeiten für Services in Publikationsangeboten steigt nach Ansicht der *Experten* durch Einbeziehung weiterer Kontextinformationen (Suchräume, semantische Analyse der Suchterme, Status der Suchenden, Verweildauer beim Lesen des Dokuments, feedbackbasierte Rankings) und qualitativen Informationen (Nutzerbewertungen der Dokumente).

Zumindest hinsichtlich der Kontextinformationen differieren die Einschätzungen aus der *Nutzerbefragung*: Für zwei abgefragte Kontextinformationen war die Zustimmung (Bewertung als häufig oder immer nützlich) gering ausgeprägt: 21,3% im Fall der Häufigkeit des Bookmarkings eines Dokuments und 15,1% für die Messung der Verweildauer der Nutzer auf einem Dokument.

Allerdings ist anzumerken, dass Nutzerbewertungen für funktionierende Services, die die genannten Kriterien etwa beim Ranking oder für Empfehlungsdienste verwenden, dennoch positiv ausfallen könnten: Die Nutzer bewerten v.a. ihnen bekannte Services, wohingegen die Experten in ihren Interviews Indikatoren für Services modellieren.

Positiver als die Kontextinformationen bewerten die Nutzer qualitative Statements über Dokumente wie Bewertungen in Form einer Notenskala (29,6% der TN sehen sie als häufig oder immer nützlich) und Nutzerbewertungen durch kurze öffentliche Kommentare (für 46,7% häufig oder immer nützlich)

Sowohl die Verwendung von Kontextinformationen und Bewertungen geht (auch wenn sie von den Experten mehrheitlich bei der Diskussion reiner Nutzungshäufigkeiten genannt wurden) über die

Anzeige rein quantitativer Daten hinaus und verweist auf Funktionen wie Rankings und Recommender die sowohl Kontextinformationen als auch Bewertungen als Faktoren einbeziehen.

Diese Perspektive nehmen auch die *Servicebetreiber* ein, die größeres Interesse an Kontextinformationen der Nutzung anmelden und teils ähnliche Informationen wie die Experten verwerten wollen, z.B. Informationen über verwendete Suchmaschinen und –begriffe oder die Aufzeichnung der Nutzungsdauer einzelner Dokumente. Neben diesen eher qualitativen Kontextinformationen haben die Betreiber allerdings auch Interesse an quantitativen Daten, die eher Monitoringbewertungen erlauben: Informationen über nationale Herkunft der Nutzer, Zeitschema der Nutzung nach Wochentagen/Uhrzeiten, Informationen über Browser/Betriebssysteme.

Recommender/ Empfehlungsdienste

Recommender werden von allen *Experten* als sinnvoll erachtet. Eine Testperson würde Recommender, die auf der Analyse sozialer Wissenschaftlernetzwerke basierten, Recommendern, die Nutzungsmuster/Zitationsmuster oder ähnliche Informationen verarbeiten, vorziehen. Überwiegend wurde aber Empfehlungsdienste als sinnvoll erachtet, die Nutzungsmuster von Dokumenten, inhaltliche Eigenschaften von Dokumenten, explizites Feedback (z.B. Bewertungen anderer Benutzer), implizites Feedback (z.B. Verweildauer auf Dokumenten), Zitationsdaten, Nutzerprofile, Nutzungsverhalten (z.B. Suchanfragen) und Autoren-Netzwerke verwenden. Teils wurde angeregt in die Analyse der Nutzungshäufigkeiten und –muster Daten aus externen Diensten miteinzubeziehen.

Diese Anregungen decken sich weitgehend mit den Ergebnissen der *Nutzerbefragung*: Recommender, die Dokumentempfehlungen auf Basis des Lese-/ Nutzungsverhaltens anbieten, wurden von 46,7% der TN als häufig oder immer nützlich eingestuft. 39,5% der TN bewerteten Recommender als häufig oder immer nützlich, die Empfehlungen auf Basis der an den Publikationen beteiligten Autoren aussprechen, und 36,4% der TN bewerteten Recommender auf Basis eines Abgleichs des eigenen Lese-/ Nutzungsverhaltens mit dem Lese-/ Nutzungsverhaltens anderer Nutzer als häufig oder immer nützlich.

Anforderungen an Recommender

Die *Experten* stellen zudem verschiedene Forderungen an Recommender. Neben einer hohen Präzision erwarten sie Transparenz der Empfehlungen. Nutzungsbasierte Zusatzfeatures können nach Meinung zweier Experten bei den Nutzern Bedenken hinsichtlich des Datenschutzes wecken, einer der Experten relativiert aber, die Qualität eines Services dürfte etwaige Datenschutzbedenken zerstreuen. Auch die *Teilnehmer der Onlinebefragung* erwarten von Empfehlungsdienste eine hohe Qualität: 75,9% der Teilnehmer bewerteten Empfehlungsdienste mit einer 90%igen Treffergenauigkeit als häufig oder immer positiv, bewerteten nur 20,1 % einen Service als häufig oder immer nützlich, der mit 50%iger Präzision passende Dokumente empfiehlt.

Verwendung der Nutzungsinformationen durch Servicebetreiber

Bei den von *Servicebetreibern* erhofften Effekten und Verwendungsszenarios standen mit Werbeeffekten, Schaffung einer Argumentationshilfe pro Open Access und Verwendung der Informationen als inneruniversitäre Argumentationshilfe Verwertungsszenarien im Vordergrund, die v.a. der eigenen Organisation bzw. dem eigenen Service dienen. Überraschend oft wurden Effekte genannt, die Monitoringcharakter haben: Performancemessung, Nutzungsanalyse und die Herstellung von Vergleichbarkeit von Repositories. In diesem Kontext lassen sich auch Punkte wie

Informationen über fehlgeschlagene Dokumentzugriffe nennen. Externe Servicefunktionen auf Basis von Nutzungsinformationen, wie die Schaffung eines Autorenservices, wurden seltener genannt. Für diese Funktionen wurden von den Betreibern entsprechende Ausgabeformen vorgesehen: Nutzungsstatistiken auf Artikel- und Autorenebene, interne Reportfunktionen (Nutzungszahlen nach Dokumenttypen, Institut, Fachbereich, Lehrstuhl etc.).

Nicht-nutzungsdatenbasierte Features

Über die genannten nutzungsdatenbasierten Features hinaus wurden auch Einschätzungen zu Funktionen abgegeben, die nicht-nutzungsdatenbasiert sind.

Die Teilnehmer der Experteninterviews nannten an dieser Stelle die vor allem Funktion, die sich auf die Nutzbarkeit von Repositories beziehen: Easy Submission, Automatisierung, Vernetzung auf Dokumentenebene, Look-and-Feel analog kommerzieller Datenbanken, intuitive Nutzungsmöglichkeiten, Nutzbarkeit von Repositories mit Mobile Devices und die intuitive Visualisierung des Repository-Contents mit Cloud Navigation.

Zusätzlich gefordert wurden die Publikationsmöglichkeiten von Preprints (Open Review), Normierung von Daten, die Verbindung von Repositories mit zugangsgeprüften Social Communities für Wissenschaftler und anpassbare Rankingoptionen.

Einen recht breiten Raum nahmen Personalisierungsoptionen ein. Diese sind vor allem auch im Hinblick auf die Recommender interessant, denn zum einen können sie wichtige Informationen für Empfehlungen liefern, zum anderen könnten Nutzer selbst bestimmen, welche Kriterien ein Recommender bei Dokumentempfehlungen berücksichtigt.